DOCUMENT RESUME

ED 395 994 TM 025 190

AUTHOR Arnold, Margery E.
TITLE The Relationship of Canonical Correlation Analysis to Other Parametric Methods.
PUB DATE 27 Jan 96
NOTE 23p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (New Orleans, LA, January 1996).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Computer Uses in Education; *Correlation; *Statistical Analysis; *Statistics; *Teaching Methods
IDENTIFIERS *Linear Models; *Parametric Analysis; Statistical Package for the Social Sciences

ABSTRACT

This paper demonstrates some characteristics that all parametric statistics have in common, using the general linear model as a framework and emphasizing that canonical correlation analysis is the general linear model. Four characteristics of the general linear model are delineated and explained using a program for the Statistical Package for the Social Sciences (SPSS). This program can be duplicated by statistics students and used to teach the concept of the general linear model. More importantly, the characteristics of the general linear model can be used to teach students about the interrelationships between all parametric statistics. Students taught this way can integrate the information about the specific models into the framework of the general model. In this manner, the learning of statistics becomes less of a chore of memorizing confusion and more of an adventure in learning and thinking. An appendix presents a program for the SPSS. (Contains 16 tables and 14 references.) (Author/SLD)

The Relationship of Canonical Correlation Analysis to Other Parametric Methods

Margery E. Arnold
Texas A&M University
College Station, TX 77843-4225

Paper presented at the annual meeting of the Southwestern Educational Research Association. January 27, 1996, New Orleans, LA.

Abstract

The present paper demonstrates some characteristics that all parametric statistics have in common, using the general linear model as a framework and emphasizing that canonical correlation analysis is the general linear model. Four characteristics of the general linear model are delineated and explained using a program for SPSS. This program can be duplicated by statistics students and used to teach the concept of the general linear model. More importantly, the characteristics of the general linear model can be used to teach students about the interrelationships between all parametric statistics. Students taught this way can integrate the information about the specific models into the framework of the general model. In this manner, the learning of statistics becomes less of a chore of memorizing confusion and more of an adventure in learning and thinking.

Several Scholars have noted that canonical correlation analysis subsumes all parametric statistics (Fan, 1992; Friedrich, 1992; Gittins, 1985, Knapp, 1978; Thompson, 1985, April). Authors have suggested that statistics be taught in the framework of the general linear model, teaching canonical correlation analysis as the final analysis and demonstrating to students how they can use canonical correlation analysis to compute all parametric statistics (cf., Friedrich, 1992; Knapp, 1978). The present paper supports that suggestion by arguing that teaching univariate and multivariate statistics using the general linear model helps students to integrate statistical concepts, as against memorizing key points that can be meaningless without context. In this model students learn the inter-relatedness of the parametric statistics (e.g., correlation, t-test, ANOVA, MANOVA, multiple regression, discriminant analysis, and canonical correlation analysis). Students integrate the information more easily when they have a framework that holds true for each analysis. The present paper outlines a framework used by Thompson (1985) and explains the basics of canonical correlation analysis and demonstrates how all parametric statistics can be calculated using the framework and canonical correlation analysis.

The General Linear Model Framework

Some students learn the framework of the general linear model in each of their statistics courses. Such students frequently hear,

1) All analyses are correlational and yield a measure of effect size that is analogous to r^2.

2) All parametric techniques invoke least-squares weights.

3) The general linear model can do anything that the specific models can do.

4) Canonical correlation analysis is the general linear model.

Line (1) of the framework refers to the concepts that links each of the specific parametric procedures to the general linear model. As Morrison (1976) explained, there are two notions that enable the link. First, one can use dummy variables to account for the degrees of freedom associated with the partitioning of a sample into nominal categories. (Canonical correlation analysis can be computed using nominal data. *Dummy variables* and *planned contrast variables* are used to compute statistics from categorical data, such as ANOVA or MANOVA, using canonical correlation analysis. This will be demonstrated later.) Secondly, differences between sample means and correlations between variables are analogous concepts. That is, a small difference between the means of two different samples on the same variable conveys the same information as a small correlation between the scores on the variable and the dummy variable of sample membership.

A demonstration of this concept helps the uninitiated reader. Gittins (1985) defines *dummy variables* as "variables whose realizations consist of arbitrary values to which meanings are assigned". (p. 68) A heuristic example of the comparison between self-esteem and pet ownership demonstrates the *dummy variable* concept. Table 1 lists hypothetical data regarding 10 children's self-esteem scores on a scale from 1 to 100 and also relates whether or not the children have a pet. Children who have a pet are given the value "1" for the dummy variable; children who do not have a pet are given the value "0" for the dummy variable. This number assignment is arbitrary. The researcher decides which group receives the "1" and which group receives the "0". Below the table are calculated (a) the difference between the mean self-esteem scores of the sample of children who have pets and the sample of children who do not have pets, and (b) the correlation between self-esteem scores and the dummy variable of pet ownership. Both statistics--the

difference between means ($\underline{\mu}_1$-$\underline{\mu}_0$– 0.6) and the correlation coefficient ($\underline{r}$=.09)--indicate that there is very little relationship between self-esteem and pet ownership.

Table 2 is provided for comparison. In this analysis there is a substantial difference between the self-esteem of pet owners and the self-esteem of non-owners. This difference is reflected in both statistics--the difference between sample means ($\underline{\mu}_1$-$\underline{\mu}_0$= 36.6) and the correlation coefficient ($\underline{r}$=.96). These two small data sets demonstrate how differences between means are related to the correlation coefficient, thus revealing the link between specific parametric procedures (such as ANOVA) and the general model (canonical correlation analysis). Furthermore, since all analyses have correlation as their basis, it is possible to calculate an $\underline{r}^2$ type of effect size in all analyses. (This point will be elaborated upon in later examples.)

The second point of the framework states that all analyses invoke least-squares weights. These weights are employed to maximize the explained variance (the correlation coefficient between the latent variable, "Yhat", and the independent variables) and to minimize the error variance. These weights are known by many names: "beta weights" in multiple regression, "discriminant function coefficients" in discriminant function analysis and MANOVA, "factor pattern coefficients" in factor analysis and "canonical function coefficients" in canonical correlation analysis. The unsuspecting graduate student might assume that these weights are all different since they all have different names. However, the weights listed above are all equivalent after a change in metric (e.g., Thompson & Borrello, 1985; Thompson, 1988). As Thompson (1991) notes,

> It is difficult to fathom why the equivalent weights used in the various parametric methods are given different names since the primary result

is confusion and the illusion that the parametric methods are different. (p. 83).

All of these weights share common properties, some of which Gittins (1985) has summarized and are listed below. Note that Gittins uses yet another name for these weights--"standardized weights". This name is probably the most sensible of all the names because it does not promote the illusion that the parametric analyses are different; rather the name emphasizes a characteristic that is common to all the weights listed above. All of the weights are *standardized* within their respective equations, and thus can be compared directly.

(a) Standardized weights are scale free.

(b) The magnitude and sign of canonical [standardized] weights can in principle at least be used as an indication of the presence of certain variables or effects and of their direction.

(c) The numerical value of canonical [standardized] weights depends on the selection of variables as well as on their scales. Addition or deletion of variables in either set is likely to produce major alterations in the remaining coefficients. Prior standardization of the observed variables to zero mean and unit variance will remove the scaling effects but the inter-dependencies will remain.

(d) Vulnerability to suppression phenomena. Suppression refers to the effect of a predictor variable which is itself positively correlated with a response variable y but which nevertheless receives a negative weight in the canonical relationship. . . .

(e) The weights tend to be highly unstable in replicate samples drawn from the same population. Several factors may

contribute to instability, notably measurement errors in the observations, collinearity of the variables of either set and inadequate sample size (Gittins, 1985, p. 25).

The third point of the framework from which to teach parametric statistics states that the general model can do anything that the specific models can do, but the specific models cannot do what the general model can do. And finally, the fourth point states that canonical correlation analysis *is* the general linear model. These points are best explained through a demonstration using SPSS. The data for this demonstration can be found in Table 3 and the SPSS commands can be found in the appendix.

The variables to be used in this demonstration include three continuous variables, Y, X, and A, two categorical variables, B and APRIME and five planned contrast variables, A1, A2, B1, A1B1 and A2B1. Variable A' or APRIME represents the partitioning of variable A into three categorical variables. Notice that any value on A of 1-4 equals a value of 1 on APRIME; whereas 5-8 and 9-12 on A equal 2 and 3, respectively, on APRIME.

The A1 variable is a planned contrast variable derived from the APRIME variable. Notice how any value of 1 on APRIME obtains a value of -1 on the A1 variable, while any value of 2 on the APRIME variable obtains a value of 1 on the A1 variable and APRIME of 3 equals A1 of 0. The A1 variable is a planned contrast variable used to compare the members of the APRIME group 1 to the APRIME group 2. APRIME group 3 is not included in this planned contrast. Following this same pattern, A2 is a planned contrast variable that compares the members of the APRIME groups 1 and 2 to the APRIME group 3 and B2 is a planned contrast variable that compares members of group 1 on variable B to members of group 0 on variable B. The

final two variables are planned contrast variables that investigate the interaction effects between A1 and B1 and between A2 and B1.

Hotelling (1936) introduced canonical correlation analysis as a way of investigating the linear correlation between two sets of continuous variables. Fisher (1936) demonstrated that canonical could be used with categorical data for the two-group case (e. g., variable B in Table 3). Rao (1948, 1952) later demonstrated that canonical correlation analysis could also be used with categorical data containing more than two groups (e. g., variable APRIME in Table 3). The remaining tables are abridged outputs from the appended computer program. The equality of results (e.g., p values and r or squared r values) establishes that canonical is, indeed, a general linear model.

Table 4 displays the results of two analyses. First, a t-test was performed using the continuous variable "Y" and the categorical variable "B". Next, a canonical correlation analysis was run using the same variables. Note that the obtained p values are exactly the same (p=.215). The obtained value of t=1.32 which when squared (t^2=1.75) exactly equals the F (F=1.75). The t was a result of the t-test analysis while the F came from the canonical analysis of the same data. You may recall that the squared t statistic with n degrees of freedom is equivalent to the F statistic with 1 and n degrees of freedom (Glass and Stanley, 1970; cited in Friedrich 1992).

Thompson (1984) has written that canonical correlation is basically a bivariate correlation coefficient.

> Conventional canonical correlation analysis investigates the degree of relationship between two sets of variables. In effect, the analysis proceeds by initially collapsing each person's scores on the variables in each variable set into a single composite variable. The simple or

bivariate correlation between the composite scores (one for each of the two variable sets) is a canonical correlation. (p. 14)

It is not surprising, then, that canonical correlation analysis can be used to compute a bivariate correlation coefficient. Table 5 displays the results of such an analysis with the results of a Pearson $\underline{r}$ correlation; both analyses were run on the same data. Note that the results are equivalent, except for the number of decimal places that are reported in the output.

Table 6 reports the results of a one-way ANOVA computed using two different procedures on the variables "Y" and "APRIME". One analysis uses the ANOVA procedure, while a second analysis used canonical correlation analysis. Table 6 shows how canonical correlation analysis can result in the same values as the ANOVA procedure. The squared canonical correlation ($\underline{R}_C$=.392) is exactly the same as eta^2 (eta^2 = SOS between/ SOS total=56/143=.3916). As in all other tables in this demonstration, the $\underline{p}$ value obtained using canonical ($\underline{p}$=.107) equals the $\underline{p}$ value obtained with the other parametric statistical procedure (in this case ANOVA, $\underline{p}$=.107).

Canonical analysis can also be used to achieve factorial ANOVA results. To compute factorial ANOVA using canonical correlation analysis it is necessary to run several canonical analyses using planned contrast variables and to do some computations. Table 7 outlines a conventional 3-by-2 ANOVA summary table for the ANOVA that was run using the dependent variable Y and two main effect categorical variables.

Table 8 details the lambdas and models used for four canonical correlation analyses. The first model used planned contrast variables for the APRIME way, the B way and the interaction effects. In this manner, it is an omnibus analysis because it tests all hypotheses. The remaining three canonical models tested all hypotheses except one. In the second model, the

APRIME hypotheses were not run. In the third model the B-way main effects were not run. In the third model, the interaction effects were not run. The analyses were done this way so that models 2, 3 & 4 could be used as divisors of the omnibus hypothesis to derive the main effect lambdas for APRIME, B and for the interaction effects. Table 9 shows these calculations. Finally, Table 10 uses a formula detailed by Rao (1952) that converts lambda (λ) to F.

$$[(1-\lambda)/\lambda] \times (\text{df error}/\ \text{df effect}) = F$$

This table shows that the Fs derived using canonical correlation analysis and the subsequent conversion are identical to the Fs derived using the factorial ANOVA procedure that were shown in Table 7.

Canonical correlation analysis is the correlation between two groups of variables (Thompson, 1984). It is not surprising then, that canonical correlation analysis can be used to compute multiple regression analyses. Table 11 displays the results of a multiple regression analysis and a canonical correlation analysis--both of which correlate the criterion variable Y with the predictor variables X, A and B. Notice that the multiple R exactly equals the canonical correlation ($\underline{R}=\underline{R}_C=.699$). The canonical lambda has been converted to the regression F, showing that the two values are equivalent. And, as in all of the other examples, the p value obtained using canonical equals the p value obtained using the parametric procedure--in this case, multiple regression. Tables 12 and 13 show the relationship between the standardized weights in canonical and those in regression. The function coefficient times the canonical correlation ($\underline{R}_C$) equals the regression beta weight ($\underline{\beta}$). Conversely, the regression beta weight ($\underline{\beta}$) divided by the canonical correlation ($\underline{R}_C$) equals the canonical function coefficient.

The final three tables in this paper demonstrate that results of multivariate statistics such as MANOVA can be obtained using the canonical

correlation procedure. These tables are analogous to the factorial ANOVA tables discussed previously. Table 14 is a MANOVA summary table while Table 15 shows the lambdas obtained using planned contrast variables and canonical correlation analysis. Table 16 shows the conversion of the canonical lambdas into MANOVA lambdas. The lambdas displayed in Table 16, derived using canonical, exactly equal the lambdas found in Table 14 which were derived using MANOVA.

Conclusion

The object of the present paper was to demonstrate the characteristics of all parametric statistics according to the general linear model while demonstrating the interrelationships between the parametric statistics and emphasizing that canonical correlation analysis is the general linear model. Four characteristics of the general linear model were delineated and explained using a program for SPSS. This program can be duplicated by statistics students and used to teach the concept of the general linear model. More importantly, the characteristics of the general linear model can be used to teach students about the interrelationships between all parametric statistics. Students taught this way can integrate the information about the specific models into the framework of the general model. In this manner, the learning of statistics becomes less of a chore of memorizing confusion and more of an adventure in learning and thinking. As Gittins (1985) has noted,

> The multivariate general linear model generates a family of methods of which canonical analysis is the most versatile. Several benefits follow from awareness of the unifying capacity of canonical analysis. Perhaps the most useful advantage is that comprehension and appreciation of a large number of statistical tools is facilitated. (p. 123)

References

Fan, Xitao. (1992, April). Canonical correlation analysis as a general analytic model. Paper presented at the annual meeting of the American Educational Research Association, San Francisco. (ERIC Reproduction Service No. ED 348 383)

Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. Annals of Eugenics, 7, 179-188.

Friedrich, K. R. (1992, January). Canonical correlation analysis: An instructional tool for all parametric statistical procedures. Paper presented at the annual meeting of the Southwest Educational Research Association, Houston, TX . (ERIC Document Reproduction Service No. 341 730)

Gittins, R. (1985). Canonical Analysis: A review with applications in ecology. Berlin: Springer-Verlag.

Glass, G. V. & Stanley, J. C. (1970). Statistical methods in education and psychology. Englewood Cliff: Prentice Hall.

Knapp, T. R. (1978). Canonical correlation analysis: A general parametric significance-testing system. Psychological Bulletin, 85, (2), 410-416.

Morrison, D. F. (1976). Multivariate Statistical Methods, 2nd ed. McGraw-Hill, New York.

Rao, C. R. (1948). The utilization of multiple measurements in problems of biological classification. Journal of the Royal Statistical Society B, 10, 159-203.

Rao, C. R. (1952). Advanced statistical methods in biometric research. New York: Wiley.

Thompson, B. (1984). Canonical correlation analysis: Uses and interpretation. Newbury Park: Sage.

Thompson, B. (1985, April). Heuristics for teaching multivariate general linear model concepts. Paper presented at the annual meeting of the American Educational Research Association, Chicago. (ERIC Reproduction Service No. ED 262 073)

Thompson, B. (1988, April). Canonical correlation analysis: An explanation with comments on correct practice. Paper presented at the annual meeting of the American Educational Research Association, New Orleans. (ERIC Document Reproduction Service No. ED 295 957)

Thompson, B. (1991). A primer on the logic and use of canonical correlation analysis. Measurement and Evaluation in Counseling and Development, 24, 80-95.

Thompson, B. & Borello, G. M. (1985). The importance of structure coefficients in regression research. Educational and Psychological Measurement, 45, 203-209.

Table 1

Heuristic Example Using a Dummy Variable to show the Link between Correlation Coefficients and Differences between Means

	Continuous Variable	Categorical Variable
Name	Self-esteem Score	Pet-ownership status
Mary	98	1
Joe	90	1
Chris	96	1
Molly	93	1
Kevin	91	1
Stacy	89	0
Matt	88	0
Sara	97	0
George	92	0
Virginia	99	0

(a) Mean for sample of children with pets = 93.6
Mean for sample of children w/out pets= 93.0
Difference between sample means= **0.6**

(b) Correlation coefficient that explains the relationship between self-esteem and the dummy variable of pet ownership (or sample membership) **r=.09**

Table 2

Heuristic Example Using a Dummy Variable to show the Link between Correlation Coefficients and Differences between Means

Name	Self-esteem Score	Pet-ownership status
Mary	98	1
Joe	90	1
Chris	96	1
Molly	93	1
Kevin	91	1
Stacy	79	0
Matt	68	0
Sara	57	0
George	42	0
Virginia	39	0

(a) Mean for sample of children with pets = 93.6
Mean for sample of children w/out pets= 57.0
Difference between sample means= **36.6**

(b) Correlation coefficient that explains the relationship between self-esteem and the dummy variable of pet ownership (or sample membership) $r = .96$

Table 3

Data used in Heuristics to Explain that Canonical is the General Linear Model

Y	X	A	B	A'	A1	A2	B1	A1B1	A2B1
1	11	5	1	2	1	-1	-1	-1	1
2	5	3	1	1	-1	-1	-1	1	1
3	2	2	1	1	-1	-1	-1	1	1
4	8	8	0	2	1	-1	1	1	-1
5	4	4	0	1	-1	-1	1	-1	-1
6	12	10	1	3	0	2	-1	0	-2
7	7	6	1	2	1	-1	-1	-1	1
8	1	1	0	1	-1	-1	1	-1	-1
9	9	12	0	3	0	2	1	0	2
10	3	7	0	2	1	-1	1	1	-1
11	6	9	0	3	0	2	1	0	2
12	10	11	1	3	0	2	-1	0	-2

Table 4

Canonical Correlation Analysis subsumes t-tests

[Interval Variable "Y" by Categorical Variable "B" (0,1)]

Canonical Analysis		t-test Analysis	
R_c*	.386		
Squared R_c	.149		
		$\underline{t}$	1.32
F	1.75	$\underline{t}^2$	1.75
$\underline{p}$	.215	$\underline{p}$	.215

*R_c is the Canonical Correlation

Table 5

Canonical Correlation Analysis Subsumes Pearson r

[Continuous Variable Y with Continuous Variable A]

Canonical Analysis		Correlation Analysis	
R_C	.566	r	.5664
Squared R_C	.321	r^2	.3208
F	4.72		
p	.055	p	.055

Table 6

Canonical Correlation Analysis Subsumes One-Way ANOVA

[Continuous Variable Y by Categorical Variable APRIME]

Canonical Analysis		ANOVA Analysis	
R_C*	.626	SOS between	56.0000
		SOS Total	143.0000
Squared R_C	.392	eta ^{2}	56/143=.3916
lambda	.60839		
F	2.89655	F	2.8966
df hypothesis	2.00	df between	2
df error	9.00	df within	9
p	.1069	p	.107

Note. The relationship between lambda and F is described in the following equation.

$[(1-\lambda)/\lambda] \times (\text{df error}/\text{ df effect}) = F$

Table 7

Factorial ANOVA: Continuous Variable "Y" by Categorical Variables "APRIME" and "B"

Source	SOS	df	MS	F	p
APRIME	56.000	2	28.000	2.754	.142
B	21.333	1	21.333	2.098	.198
APRIME B	4.667	2	2.333	.230	.802
Error	61.000	6	10.167		
Total	143.000	11	13.000		
eta^2	82/143=.573				

Table 8
Canonical Analysis Using Four Models

Model	Predictors of Y	Lambda
1: Omnibus	A1, A2, B1, A1B1, A2B1	.42657
2: No APRIME	B1, A1B1, A2B1	.81818
3: No B	A1, A2, A1B1, A2B1	.57576
4: No interaction	A1, A2, B1	.45921

Table 9
Calculating Lambda for each Source of Variance

Source	Models	Calculation	Lambda
APRIME	Model 1/Model 2	.42657/.81818	.52136
B	Model 1/Model 3	.42657/.57576	.74088
APRIME B	Model 1/Model 4	.42657/.45921	.92892

Table 10

Conversion of Canonical's Lambdas to ANOVA'S Fs

Source	[(1-lambda)/lambda] * (df error/df effect)	= F
APRIME	[(1-.52136)/.53136]*(6/2)	
	.91806 * 3	= 2.754
B	[(1-.74088)/.74088]*(6/1)	
	.34975*6	= 2.098
APRIME B	[(1-.92892)/.92892]*(6/2)	
	.07652 * 3	= .230

Table 11

Canonical Correlation Subsumes Multiple Regression

Variable "Y" with Variables X, A & B

Canonical Analysis		Regression Analysis	
R_C	.699	R	.69920
Squared R_C	.836	Squared R	.83618
Lambda	.30080		
Conversion to F	[(1-.3008)/.3008][8/3]	F	6.19861
	=6.1985816		
p	.018	p	.0175

Table 12

Canonical Function Coefficients Converted to Regression Beta Weights

Predictor	Function Coefficients	(Canonical Correlation)	(Beta Weights)
X	(-1.18685)	(.83618)=	-.992424
A	(1.54635)	(.83618)=	1.293030
B	(.14582)	(.83618)=	.121930

Table 13

Regression Beta Weights Converted to Canonical Function Coefficients

Predictor	(Beta Weights)	/(Canonical Correlation)	= Function Coefficients
X	(-.992424)	/(.83618)=	-1.18685
A	(1.293030)	/(.83618)=	1.54635
B	(.121930)	/(.83618)=	.14582

Table 14

Factorial MANOVA: "Y","X" with "APRIME" and "B"

Source	SOS	df	p
APRIME	.03202	4, 10	.001
B	.60902	2, 5	.289
APRIME B	.37812	4, 10	.257

Table 15
Canonical Analysis Using Four Models

Model	Predictors of Y, X	Canonical Lambda
1: Omnibus	A1, A2, B1, A1B1, A2B1	.02113
2: No APRIME	B1, A1B1, A2B1	.65989
3: No B	A1, A2, A1B1, A2B1	.03469
4: No interaction	A1, A2, B1	.05588

Table 16
Canonical's Lambda for each Source of Variance Equals MANOVA's Lambda

Source	Models	Calculation	Lambda
APRIME	Model 1/Model 2	.02113/.65989	.03202
B	Model 1/Model 3	.02113/.03469	.60911
APRIME B	Model 1/Mod[illegible]	.02113/.05598	.03781

Appendix

```
TITLE '$$$ Show how canonical subsumes other methods' .
DATA LIST FILE='A:GLM.DTA'/1
  Y 1-2 X 4-5 A 7-8 B 10 APRIME 12 A1 14-15 A2 17-18 B1 20-21
  A1B1 23-24 A2B1 26-27 .
LIST VARIABLES=ALL/CASES=50/FORMAT=NUMBERED .
SUBTITLE '1.  CCA subsumes t-tests #######' .
T-TEST GROUPS=B(0,1)/VARIABLES=Y .
MANOVA Y WITH B/PRINT=SIGNIF(MULTIV EIGEN DIMENR)/
  DISCRIM=STAN CORR ALPHA(.99)/DESIGN .
SUBTITLE '2.  CCA subsumes Pearson r #######' .
CORRELATIONS VARIABLES=Y A .
MANOVA Y WITH A/PRINT=SIGNIF(MULTIV EIGEN DIMENR)/
  DISCRIM=STAN CORR ALPHA(.99)/DESIGN .
SUBTITLE '3.  CCA subsumes ANOVA #######' .
ONEWAY Y BY APRIME(1,3) .
MANOVA A1 A2 WITH Y/PRINT=SIGNIF(MULTIV EIGEN DIMENR)/
  DISCRIM=STAN CORR ALPHA(.99)/DESIGN .
SUBTITLE '4.  CCA subsumes one-way ANOVA #######' .
ONEWAY Y BY APRIME(1,3) .
MANOVA A1 A2 WITH Y/PRINT=SIGNIF(MULTIV EIGEN DIMENR)/
  DISCRIM=STAN CORR ALPHA(.99)/DESIGN .
SUBTITLE '5.  CCA subsumes factorial ANOVA #######' .
ANOVA Y BY APRIME(1,3) B(0,1) .
MANOVA A1 A2 B1 A1B1 A2B1 WITH Y/PRINT=SIGNIF(MULTIV EIGEN DIMENR)/
  DISCRIM=STAN CORR ALPHA(.99)/DESIGN .
MANOVA B1 A1B1 A2B1 WITH Y/PRINT=SIGNIF(MULTIV EIGEN DIMENR)/
  DISCRIM=STAN CORR ALPHA(.99)/DESIGN .
MANOVA A1 A2 A1B1 A2B1 WITH Y/PRINT=SIGNIF(MULTIV EIGEN DIMENR)/
  DISCRIM=STAN CORR ALPHA(.99)/DESIGN .
MANOVA A1 A2 B1 WITH Y/PRINT=SIGNIF(MULTIV EIGEN DIMENR)/
  DISCRIM=STAN CORR ALPHA(.99)/DESIGN .
SUBTITLE '6.  CCA subsumes multiple R #######' .
REGRESSION VARIABLES= Y X A B/DEPENDENT=Y/ENTER X A B .
MANOVA X A B WITH Y/PRINT=SIGNIF(MULTIV EIGEN DIMENR)/
  DISCRIM=STAN CORR ALPHA(.99)/DESIGN .
SUBTITLE '7.  CCA subsumes factorial MANOVA #######' .
MANOVA Y X BY APRIME(1,3) B(0,1)/PRINT=SIGNIF(MULTIV EIGEN DIMENR)/
  DESIGN .
MANOVA  A1  A2  B1  A1B1  A2B1  WITH  Y  X/PRINT=SIGNIF(MULTIV  EIGEN
DIMENR)/
  DISCRIM=STAN CORR ALPHA(.99)/DESIGN .
MANOVA B1 A1B1 A2B1 WITH Y X/PRINT=SIGNIF(MULTIV EIGEN DIMENR)/
  DISCRIM=STAN CORR ALPHA(.99)/DESIGN .
MANOVA A1 A2 A1B1 A2B1 WITH Y X/PRINT=SIGNIF(MULTIV EIGEN DIMENR)/
  DISCRIM=STAN CORR ALPHA(.99)/DESIGN .
MANOVA A1 A2 B1 WITH Y X/PRINT=SIGNIF(MULTIV EIGEN DIMENR)/
  DISCRIM=STAN CORR ALPHA(.99)/DESIGN .
```